

Algorithm for Explaining Credit Scores

RELATION TO OTHER PATENT APPLICATIONS

5

This document derives priority from provisional U.S. patent application serial no. 60/222,231, filed 8/1/00; and is related to U.S. patent application serial no. 09/790,453, filed 2/22/01, which in turn claims benefit of provisional U.S. patent application serial no. 60/222,205, filed 8/1/00.

10

BACKGROUND OF THE INVENTION

TECHNICAL FIELD

15

The invention relates to credit scoring. More particularly, the invention relates to an algorithm for explaining credit scores.

DESCRIPTION OF THE PRIOR ART

20

Recent events have made it desirable for developers of credit scoring algorithms, such as Fair, Isaac and Company, Inc. of San Rafael, California (FICO) to move toward offering a service to deliver credit bureau risk scores and explanations directly to consumers and lenders. Consumer advocacy groups and credit counseling organizations have provided positive feedback on these announced intentions. Additionally, credit scoring developers clients, *i.e.* the credit grantors themselves, have expressed their understanding of the need to pursue this undertaking. Most organizations

25

are comfortable that each credit scoring developer, such as Fair, Isaac, is the only entity in the market that can actively take on the role of credit score delivery and explanation.

5 A comprehensive score delivery and explanation service should include all of the following pieces:

1. Credit scores delivered to consumers.

10 2. The primary reason codes that describe why the score was not higher.

3. The consumer's credit bureau report from which the score was calculated to allow them to cross-reference the information with his/her actual credit report.

15 4. A personalized score explanation that describes to that consumer, in plain language, how their individual score was derived. This explanation service can be further enhanced using data elements present in the consumer's credit report.

20 Given the desirability of providing such information to consumers, it would be advantageous to provide a method and apparatus for explaining credit scores.

A. Flint, D. Lear, C. St. John, *Method and Apparatus for Explaining Credit*
25 *Scores*, U.S. patent application serial no. 09/790,453 (2/22/01) describe a Web site containing an array of informative resources including for-pay

services and extranet functions to serve consumers and traditional players in the financial services industry, including financial counselors, mortgage brokers, direct lenders, large national credit issuers, and third-party credit report re-sellers, plus information seekers such as the press, consumer groups, and government agencies. A primary focus of the *Flint et al.* invention is to educate consumers, consumer groups, and the consumer press by offering them access to the exceptionally high-quality information, both general and personal, about the practices of collection, storing, reporting, and evaluating consumer credit data.

It would be advantageous to provide an algorithm for explaining credit scores, for example in connection with a credit score explanation service.

SUMMARY OF THE INVENTION

An exemplary Web-based score explanation service typically requires only the credit bureau identifier, credit score, and up to four reason codes as input. The invention herein discloses an algorithm that is used to provide an explanation of the primary factors influencing the score based upon a rich data feed.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block schematic diagram showing targeted users, access and entry points, and services provided by myFICO.com; and

Fig. 2 is a flow diagram showing an algorithm for explaining credit scores according to the invention.

DETAILED DESCRIPTION OF THE INVENTION

5

The presently preferred embodiment of the herein described algorithm for explaining credit scores is provided for use in conjunction with a credit score explanation service that may be implemented in any of several embodiments.

The preferred embodiment of the invention operates in conjunction with a

10 Web site containing an array of informative resources including for-pay services and extranet functions to serve consumers and traditional players in the financial services industry, including financial counselors, mortgage brokers, direct lenders, large national credit issuers, and third-party credit report re-sellers, plus information seekers such as the press, consumer
15 groups, and government agencies. A primary focus of such Web site is to educate consumers, consumer groups, and the consumer press by offering them access to the exceptionally high-quality information, both general and personal, about the practices of collection, storing, reporting, and evaluating consumer credit data.

20

The working title of the Fair, Isaac and Company (FICO) Web presence is myFICO (myfico.com) because the most visible elements of the service are aimed directly at consumers who want to learn about their own FICO score.

25 Although the on-demand receipt of FICO scores is thought to be the primary draw to the site (based on consumer interest and press coverage), the

invention also offers access to additional valuable services, such as registration in an opt-in/opt-out database, the ability to initiate requests for credit investigations, the ability to link to consumer credit counseling services should scores be low and represent high risk, and the ability to access multiple reports from different repositories upon request. These services heighten the level of consumer education, and also offer individuals access to information, actions, and preferences they have not had previously.

An additional benefit is to use myFICO.com to supply the consumer with their score and if that score is sufficient to pass the cutoff scores of specific brokers or lenders, the credit scoring developer can pass the consumer's name, application, and credit score on to the lender for consideration. The invention allows the credit scoring developer to build broker networks to refer these applicants to lenders who would approve them. The credit scoring developer can also link the applicants' email address to credit companies who wish to pre-approve and solicit these consumers based on score. This is a much more cost effective origination process (via email) than direct mail today.

Fig. 1 is a block schematic diagram showing targeted users, access and entry points, and services provided by myFICO.com. Access by consumers may be through a credit reporting agency, using an identification verification process to access credit score reports, for opt-in/opt-out requests, to access a report service, and to initiated on line investigations; through a secure, one time connection for one-off credit score reports; or through an entirely anonymous access method (the latter also allows access by government agencies for consumer oriented

information 21. myFICO.com also provides an extranet logon facility 25 to credit score reports 37 for such users as financial counselors 24, mortgage brokers 26, and direct lenders 27; an automated application service provider entry 29 to credit score reports 30 and other reports 31 for such users as large credit issuers 28, on line financial service providers 32, and credit report resellers 33; and repository access 35 to credit score reports and other reports 36 for repository consumer representatives 34 and credit report resellers 33. See A. Flint, D. Lear, C. St. John, *Method and Apparatus for Explaining Credit Scores*, U.S. patent application serial no. 09/790,453 (2/22/01).

Score Explanation Service

An exemplary Web-based score explanation service requires, for example, only the credit bureau identifier, credit score, and up to four reason codes as input. In contrast, the invention herein discloses an algorithm that is used to provide an explanation of the primary factors influencing the score based upon a rich data feed. This algorithm can be enhanced depending upon the amount of input data available, although the use of an enhanced algorithm is optional and not considered to be a key element of the subject invention. The actual explanations may be selected as appropriate for the application to which the invention is put. Typical explanations are those described in A. Flint, D. Lear, C. St. John, *Method and Apparatus for Explaining Credit Scores*, U.S. patent application serial no. 09/790,453 (2/22/01).

Fig. 2 is a flow diagram showing an algorithm for explaining credit scores

according to the invention. The invention comprises a score explanation algorithm, which could be applied to any score. The presently preferred embodiment of the invention provides the basis for a general web-based score explanation service (see Fig. 1).

5

Score Explainer

Consider a score, which can be written as some function of a set of prediction characteristics (100)

10

$$\begin{aligned} \text{Score} &= \varsigma(\chi_1, \chi_2, \dots, \chi_c), \\ \text{where} \\ \chi_j &= \text{Prediction Characteristic } j. \end{aligned}$$

For a credit bureau risk score, there are about 80 prediction characteristics, which can include, for example, such characteristics as the number of trade lines with a current delinquency, although the actual number chosen when practicing the herein disclosed invention can vary depending upon the particular application to which the invention is put. If the score is a single scorecard, *i.e.* a scoring model where the score for an individual is the sum of their characteristic scores, and where the number of terms in this sum is the number of characteristics, so that, for each characteristic, the individual is assigned a score weight and then their final score is the sum these score weights, then it can be written in the form

$$\text{Score} = \sum_{j=1}^c \varsigma_j(\chi_j).$$

If the score is a segmented scorecard, *i.e.* where the population is segmented into mutually exclusive segments, and where a separate scorecard model is developed for each segment, then the formula is more complicated. For example the score weight associated with a particular value of a particular characteristic depends on what segment the individual is in. In such cases, the formula can be written down and analyzed. In fact, the invention disclosed herein works for any score, which can then be computed in a reasonably fast manner by application of the invention thereto.

To explain a score in detail, we define a set of surrogate characteristics, z_1, z_2, \dots, z_p , which are labeled "Areas for improvement" (110). Areas for improvement include, for example, such factors as too much delinquency, too much debt, short credit history, etc. There are p Areas for Improvement, each area represented mathematically by a surrogate characteristic. Typically, $p < c$. For example, see below.

Use these Areas for Improvement surrogate characteristics to develop a surrogate score of the form

$$\psi(z_1, z_2, \dots, z_p),$$

which is developed using z_1, z_2, \dots, z_p as the prediction characteristics and

$$y = \varsigma(\chi_1, \chi_2, \dots, \chi_c)$$

as the performance (dependent) variable (120). Now consider the customer,

5 who wants their score explained. Their actual score is

$$\text{Score}^* = \varsigma(\chi_1^*, \chi_2^*, \dots, \chi_c^*)$$

and their values of z_1, z_2, \dots, z_p are

$$z_1^*, z_2^*, \dots, z_p^*.$$

10 A rich data feed with the values of the z 's is needed, but not necessarily the values of the x 's.

Associated with each Area for Improvement, define the potential improvement metric (130)

$$I_k = 100 \times \frac{\left[\max_{z_k} \{ \psi(z_1^*(z_k), \dots, z_k, \dots, z_p^*(z_k)) - \psi(z_1^*, \dots, z_k^*, \dots, z_p^*) \} \right]}{\text{Score}^*}$$

15

= Maximum possible percent improvement for
Area of Improvement k ,

where, e.g. $z_1^*(z_k)=z_1^*$ unless z_1^* cannot coexist with z_k . In that case, $z_1^*(z_k)=E[z_1|z_k]$, or some other value of z_1 that can coexist with z_k .

5 Suppose that the values of the I's are ordered as follows:

$$I_4 > I_7 > I_2 > I_{11} > \dots$$

Then one could optionally create an ordered Area for Improvement table of the form shown in Table “A” below (140).

10

Table A: Areas for Improvement

Areas for Improvement	Potential Percent Improvement in Score
4	I_4
7	I_7
2	I_2
11	I_{11}
:	:
:	:

15 The number of Areas for Improvement listed would be small relative to the value of p .

Definitions of the Areas for Improvement

In the simple case, one could use the original prediction characteristics as the Areas for Improvement (152). However, in cases such as the credit bureau risk score, this would yield too many Areas for Improvement.

5 Another possibility is to use the set of prediction characteristics, which are computed, for example, as part of Fair Isaac's Search™ Software product as the Areas for Improvement (154). When Search™ goes to the credit bureau to get information on a person, these Search™ prediction characteristics are returned. The Search™ Software product improves origination decision making by automatically obtaining credit bureau reports and scores from the major North American consumer credit bureaus, as well as providing comprehensive and sophisticated analysis (see <http://www.fairisaac.com/index.html>) In Search™ there are about 40 characteristics.

15 Another approach is to define a set of surrogate characteristics from some standard categorization of the credit bureau characteristics that go into scores (150). The categories listed on a Fair Isaac Web page (see <http://www.myfico.com/filfsf.html>) are (i) Payment history, (ii) Amounts Owed, (iii) Length of credit history, (iv) New credit, and (v) Types of credit use.

20 Suppose that the 80 original characteristics are categorized into the five categories above as follows:

Category 1: x_1, x_2, \dots, x_{15}

25 Category 2: $x_{16}, x_{17}, \dots, x_{35}$

Category 3: $x_{36}, x_{37}, \dots, x_{50}$

Category 4: $x_{51}, x_{52}, \dots, x_{65}$

Category 5: $x_{66}, x_{67}, \dots, x_{80}$

Then one could develop a surrogate characteristic of the form

$$z_1 = \phi_1(x_1, x_2, \dots, x_{15}),$$

where

$$\phi_1(x_1, x_2, \dots, x_{15})$$

is a score developed using x_1, x_2, \dots, x_{15} as the prediction characteristics and the real credit score as the performance (dependent) variable (160). The variables z_2, z_3, \dots, z_5 could be developed in a similar way.

- 15 This creates five possible Areas for Improvement (170), which might be too small. Each of the above categories could be broken into several sub-categories to create more Areas for Improvement.

The three approaches described above are shown on Fig. 2 as alternatives.

Discussion

Reason Codes

The following discussion of the invention is a generalization of the technique used today for computing reason codes. Associated with a person's score is a set of about four reason codes, which can be easily returned from a credit bureau along with the score value. In the background, associated with each reason code, is a score difference. These score differences are defined slightly differently than the score differences in the formula for I_k but they could be converted to percentages in the same way that the score differences are converted to percentages. This idea is readily implemented with a data feed, which includes the score differences associated with the reason codes.

Segment Split Variables

The above formula for I_k takes into account automatically and precisely the split variables used to define the segments in a segmented score. The maximization over z_k for a split variable involves the computation of the customer's score in several segments, because, as z_k varies over its range, one moves from segment to segment.

Rich Data Feed

The presently preferred embodiment of the herein disclosed invention requires a rich data feed, *i.e.* the values of the z^{**} 's. However, the values of the x^{**} 's are not needed, unless they are the same as the z^{**} 's.

Number of Areas for Improvement

In the preferred embodiment, the total number of Areas for Improvement could be just about any number between five and 80. However, showing the customer the typical top four is presently preferred. The total number is preferably significantly more than four, *i.e.* five is too small. Twenty is presently preferred. This means that the five categories mentioned above in connection with the Fair Isaac site should be expanded to about twenty. The original 80 prediction characteristics should be put into about twenty categories.

Generality

The algorithm described herein would work for any score, which can then be computed in a reasonably fast manner by application of the invention herein thereto. This includes neural networks and regression trees.

Although the invention is described herein with reference to the preferred embodiment, one skilled in the art will readily appreciate that other applications may be substituted for those set forth herein without departing from the spirit and scope of the present invention. Accordingly, the invention should only be limited by the Claims included below.